COMPLETE CLASS THEOREMS DERIVED FROM
CONDITIONAL COMPLETE CLASS THEOREMS

by

Morris L. Eaton [1/]

University of Minnesota

Technical Report No. 270

July, 1976

[1/] This research was supported in part by a grant from the National Science Foundation - NSF - GP - 34482.

Abstract

Let $(\mathfrak{X}, \mathcal{B}_1, \mu)$ and $(\mathcal{Y}, \mathcal{B}_2, \nu)$ be σ-finite measure spaces and suppose Θ is a separable metric space. Let $f(x|y, \theta)$ be a family of conditional densities on $(\mathfrak{X}, \mathcal{B}, \mu)$. Consider an action space A which is a compact metric space with $\mathcal{B}_A$ the Borel σ-algebra and a loss function $L(\theta, a)$ such that $L(\theta, \cdot)$ is continuous. For any decision rule δ: $\mathcal{B}_A \times \mathfrak{X} \to [0, 1]$, assume the risk function $R(\delta, \cdot)$ is continuous on Θ. Suppose that a set of decision rules $\mathfrak{m}_0$ is an essentially complete class for each $y \in \mathcal{Y}$ for the conditional decision problem. Let $\mathfrak{m}^*$ be the set of decision rules η: $\mathcal{B}_A \times (\mathfrak{X} \times \mathcal{Y}) \to [0, 1]$ such that $\eta(\cdot \mid \cdot, y) \in \mathfrak{m}_0$ a.e. $[\nu]$. Then $\mathfrak{m}^*$ is an essentially complete class no matter what the family of marginal densities on the space $(\mathcal{Y}, \mathcal{B}_2, \nu)$.

AMS Subject Classification: Primary 62C07

Key words and phrases: decision theory, complete class, conditional complete class.

Introduction:

The problem of obtaining complete class theorems from conditional complete class theorems is well illustrated by the following example. Suppose $f(x|y, \theta)$ is a conditional density of X given $Y = y$ where X and Y are real valued random variables and $\theta \epsilon \Theta$ - an interval of the line. Let $g(y|\theta)$ denote the marginal density of Y. Consider the problem of testing H_0: $\theta \leq \theta_0$ versus H_1: $\theta > \theta_0$ and assume $f(x|y, \theta)$ has a monotone likelihood ratio in x and θ for each fixed y. Let $\mathcal{D}$ be the class of test functions (for y fixed) of the form

$$(1) \qquad \varphi(x) = \begin{cases} 0 & \text{if } x < x_0 \\ \gamma & \phantom{\text{if } x} = \\ 1 & \phantom{\text{if } x} > \end{cases}$$

As is well known, $\mathcal{D}$ is an essentially complete class of tests for y fixed. Now, let $\eta(x, y)$ be any test function for H_0 versus H_1 based on both X and Y. For each fixed y, the essential completeness of $\mathcal{D}$ implies there is a test function $\varphi_y \epsilon \mathcal{D}$ which is at least as good as $\eta(\cdot, y)$; that is,

$$(2) \qquad \begin{cases} \mathcal{E}_\theta(\varphi_y(X)|Y = y) \leq \mathcal{E}_\theta(\eta(X, y)|Y = y), & \theta \leq \theta_0 \\ \qquad\qquad\qquad\quad \geq \qquad\qquad\qquad\qquad\quad , & \theta > \theta_0 \end{cases}$$

<u>If</u> $\varphi_Y(X)$ were a jointly measurable function of (X, Y), one could then integrate (with respect to the marginal distribution of Y) both sides of the two inequalities in (2) to obtain

$$(3)\qquad \begin{cases} \mathcal{E}_\theta \varphi_Y(X) & \leq \quad \mathcal{E}_\theta \eta(X, Y), \quad \theta \leq \theta_0 \\ & \geq \quad \qquad\qquad , \quad \theta > \theta_0. \end{cases}$$

Let $\mathcal{D}^*$ be the set of test functions $\xi(X, Y)$ such that $\xi(\cdot, y)\epsilon\mathcal{D}$ for each $y\epsilon\mathcal{D}$. Then the above argument yields (assuming the measurability of $\varphi_Y(X)$) that $\mathcal{D}^*$ is an essentially complete class. Thus, it is clear that the only difficulty with this argument is in showing that one can select a measurable version of $\varphi_y\epsilon\mathcal{D}$. It is precisely this measurability problem which arose in the work of Matthes and Truax (1967) concerning complete class results for testing problems in multivariate exponential families. That this measurability problem is fairly non-trivial is evidenced by the fact that Matthes and Truax (1967) found it necessary to use the Martingale Convergence Theorem and a version (involving measurability) of the Blaschke Selection Theorem to show the existence of a measurable $\varphi_Y(X)$.

The purpose of this paper is to establish complete class results from conditional complete class results for a fairly general decision problem with a compact action space. As with the above example and the Matthes-Truax problem, the primary difficulty is the measurability. We have found a recent result of Brown and Purves (1973) useful in this context. In Section 2, we describe the decision problem under consideration and discuss a representation theorem for decision functions given in Farrell (1967). In addition, the result of Brown and Purves (1973) is outlined. The main result of this paper is proved in Section 3. Basically, this result says that if $\mathfrak{m}$ is an essentially complete class for a conditional (given $Y = y$) decision problem, then $\mathfrak{m}^*$, the class of decision functions which are in $\mathfrak{m}$ for y fixed, is essentially complete for the unconditional decision problem.

The main result of Section 3 is not applicable to all the examples to which the Matthes-Truax result can be applied. This point, together with some applications, is discussed in Section 4. The appendix of this paper establishes some measurability results needed in Section 3.

§2 Notation and Assumptions:

This section consists primarily of a long string of definitions, notation and assumptions concerning the structure of the decision problem under study. The reader is urged to keep in mind the testing problem treated by Matthes and Truax (1967) as this problem is what motivated the current work. The following notation and assumptions hold throughout.

(2.1) $\begin{cases} (\mathfrak{X}, \mathcal{B}_1, \mu) \text{ is a } \sigma\text{-finite measure space such that} \\ L_1(\mathfrak{X}, \mathcal{B}_1, \mu) \text{ is a separable Banach space.} \end{cases}$

(2.2) $\begin{cases} (\mathfrak{Y}, \mathcal{B}_2, \nu) \text{ is a } \sigma\text{-finite measure space such that } \mathfrak{Y} \\ \text{is a complete separable metric space and } \mathcal{B}_2 \text{ is the} \\ \sigma\text{-algebra of Borel sets.} \end{cases}$

(2.3) $\begin{cases} \Theta \text{ denotes the parameter space of the decision problem and} \\ \Theta \text{ is assumed to be a separable metric space.} \end{cases}$

(2.4) $\begin{cases} f(\cdot|y, \theta),\ y\epsilon\mathfrak{Y},\ \theta\epsilon\Theta \text{ is a family of densities on} \\ (\mathfrak{X}, \mathcal{B}_1, \mu) \text{ and } f(\cdot|\cdot, \theta) \text{ is } \mathcal{B}_1\times\mathcal{B}_2 \text{ measurable for} \\ \text{each } \theta\epsilon\Theta. \end{cases}$

(2.5) $\begin{cases} A \text{ denotes the action space of the decision problem, } A \text{ is a} \\ \text{compact metric space, and } \mathcal{B}_A \text{ is the } \sigma\text{-algebra of Borel sets.} \\ C(A) \text{ denotes the Banach space of continuous functions on } A \\ \text{(with the sup-norm).} \end{cases}$

(2.6) $\begin{cases} W: \Theta \times A \to [0, \infty) \text{ is the loss function for the decision prob-} \\ \text{lem and } W(\theta, \cdot) \text{ is continuous on } A \text{ for each } \theta\in\Theta. \end{cases}$

<u>Definition:</u> A function $\delta: \mathcal{B}_A \times \mathcal{X} \to [0, 1]$ is a <u>decision</u> <u>function</u>, if

(i) $\delta(\cdot|x)$ is a probability measure on $\mathcal{B}_A$ for each $x\in\mathcal{X}$

(ii) $\delta(B|\cdot)$ is $\mathcal{B}_1$ measurable for each $B\in\mathcal{B}_A$.

(2.7) $\begin{cases} \text{If } \delta \text{ is a decision rule, let} \\ R_y(\delta, \theta) = \iint W(\theta, a)\delta(da|x)f(x|y, \theta)\mu(dx) \text{ and} \\ \text{assume } R_y(\delta, \cdot) \text{ is continuous on } \Theta \text{ for each } y \text{ and } \delta. \end{cases}$

Suppose δ_1 and δ_2 are two decision rules such that for all $g\in C(A)$ and $h\in L_1$ $(= L_1(\mathcal{X}, \mathcal{B}_1, \mu))$

(2.8) $$\iint g(a)h(x)\delta_1(da|x)\mu(dx) = \iint g(a)h(x)\delta_2(da|x)\mu(dx)$$

Then, using the separability of $C(A)$, it is not hard to show there is a μ null set, say N, such that for all $x\notin N$, $\delta_1(\cdot|x) = \delta_2(\cdot|x)$. Conversely, if $\delta_1(\cdot|x) = \delta_2(\cdot|x)$ a.e. (μ), then it is clear that (2.8) holds. For any decision rule δ, $[\delta]$ denotes the equivalence class of decision rules which are equivalent to δ as described above. Let

(2.9) $$\mathfrak{m} = \{[\delta] | \delta \text{ is a decision rule}\}.$$

We will write "$\delta\in\mathfrak{m}$" when there is no reason to distinguish between δ and $[\delta]$.

Following Farrell (1964), it is convenient to think of $\mathfrak{M}$ as a subset of the set of continuous bilinear functionals on $C(A)\times L_1$. For $\delta\epsilon\mathfrak{M}$, consider $[\cdot, \cdot]_\delta$ defined on $C(A)\times L_1$ by

$$(2.10) \qquad [g, h]_\delta \equiv \iint g(a)h(x)\delta(da|x)\mu(dx).$$

Clearly $[\cdot, \cdot]_\delta$ is bilinear on $C(A)\times L_1$ and satisfies

$$(2.11) \qquad \begin{cases} \text{(i)} & [1, h]_\delta = \int h(x)\mu(dx) \\ \text{(ii)} & g \geq 0, \quad h \geq 0 \quad \text{implies} \quad [g, h]_\delta \geq 0. \\ \text{(iii)} & \sup\limits_{\substack{\|g\|=1 \\ \|h\|=1}} [g, h]_\delta = 1 \end{cases}$$

Conversely, suppose $[\cdot, \cdot]$ is a bilinear functional on $C(A)\times L_1$ which satisfies (2.11)(i), (ii), (iii) (without the subscript δ). It follows from the Appendix in Farrell (1964) that there is a decision rule δ such that $[\cdot, \cdot] = [\cdot, \cdot]_\delta$.

To introduce a topology on $\mathfrak{M}$, for each $g\epsilon C(A)$ amd $h\epsilon L_1$, define $T_{g,h}:\mathfrak{M} \to (-\infty, \infty)$ by

$$(2.12) \qquad T_{g,h}(\delta) \equiv [g, h]_\delta.$$

The weakest topology such that all $T_{g,h}$, $g\epsilon C(A)$, $h\epsilon L_1$, are continuous is the <u>weak</u> topology on $\mathfrak{M}$. Since $C(A)$ and L_1 are both separable, it follows that this topology is metric. In addition, a standard embedding argument shows that $\mathfrak{M}$ is compact in the weak topology. In summary, under the assumptions we've made on L_1 and A, $\mathfrak{M}$ is a compact metric space with the weak topology.

This section is concluded with a statement of a result due to Brown and Purves (1973). Let $\mathcal{U}$ and $\mathcal{V}$ be metric spaces. If $E \subseteq \mathcal{U}\times\mathcal{V}$, $\text{proj}(E) \equiv \{u \mid (u,v)\in E \text{ for some } v\in\mathcal{V}\} \subseteq \mathcal{U}$.

<u>Definition:</u> $\mathcal{S} \subseteq E \subseteq \mathcal{U}\times\mathcal{V}$ is a <u>Borel selection</u> if

(i) $\mathcal{S}$ is a Borel set in $\mathcal{U}\times\mathcal{V}$

(ii) For each $u\in\mathcal{U}$, $\mathcal{S}_u \equiv \{v\in\mathcal{V} \mid (u, v)\in\mathcal{S}\}$ contains at most one point.

(iii) $\text{proj}(\mathcal{S}) = \text{proj}(E)$.

For each selection $\mathcal{S}$ is the function $\rho: \text{proj}(\mathcal{S}) \to \mathcal{V}$ which assigns to each $u\in \text{proj}(\mathcal{S})$, the unique $v\in\mathcal{V}$ such that $(u, v)\in\mathcal{S}$. Thus $(u, \rho(u))\in E$ for all $u\in \text{proj}(E)$.

<u>Theorem (Brown and Purves (1973):</u> Let $\mathcal{U}$, $\mathcal{V}$ be complete separable metric spaces and $E \subseteq \mathcal{U}\times\mathcal{V}$ be a Borel set. If for each $u\in\mathcal{U}$, the section $E_u \equiv \{v\in\mathcal{V} \mid (u, v)\in E\}$ is σ-compact, then there is a Borel selection $\mathcal{S}$. Further, proj(E) is a Borel set and ρ is a Borel measurable function defined on proj(E) (= proj($\mathcal{S}$)).

<u>§3 A Complete Class Theorem:</u>

With the notation and assumptions in Section 2, we now want to prove a complete class result for an unconditional decision problem (decision functions are functions of both X and Y) given a complete class theorem for the conditional problem (y fixed). Assume that $\mathfrak{m}_0 \subseteq \mathfrak{m}$ is an essentially complete class for the decision problem given in Section 2 for each fixed y, and $\mathfrak{m}_0$ is a closed (hence compact) subset of $\mathfrak{m}$. That is, given any decision rule $\delta\epsilon\mathfrak{m}$ and $y\epsilon\mathcal{Y}$, there is a decision rule $\delta_1\epsilon\mathfrak{m}_0$ (which can depend on y) such that

$$(3.1) \qquad R_y(\delta_1, \theta) \leq R_y(\delta, \theta) \quad \text{for all} \quad \theta\epsilon\Theta.$$

In what follows, η denotes a decision rule defined on $\mathcal{B}_A \times (\mathcal{X}\times\mathcal{Y})$ to $[0, 1]$. Thus $\eta(\cdot \mid x, y)$ is a probability measure on $\mathcal{B}_A$ and $\eta(B \mid \cdot, \cdot)$ is $\mathcal{B}_1\times\mathcal{B}_2$ measurable for each $B\epsilon\mathcal{B}_A$. Given such an η, $\eta_y\epsilon\mathfrak{m}$ is defined by $\eta_y(B\mid x) \equiv \eta(B \mid x, y)$. Define a set of decision rules $\mathfrak{m}^*$ by

$$(3.2) \qquad \mathfrak{m}^* = \left\{ \eta \,\middle|\, \begin{array}{l} \eta \text{ is a decision rule and } \{y \mid [\eta_y] \notin \mathfrak{m}_0\} \text{ is a} \\ \nu \text{ null set.} \end{array} \right\}$$

<u>Theorem 1:</u> Given any decision rule η_0, there is an $\eta\epsilon\mathfrak{m}^*$ and a ν null set N (depending on η_0 and η) such that if $y \notin N$

$$(3.3) \qquad R_y(\eta_y, \theta) \leq R_y(\eta_{0,\,y}, \theta) \quad \text{for all} \quad \theta\epsilon\Theta.$$

<u>Proof:</u> Define a set $E \subseteq \mathcal{Y}\times\mathfrak{m}_0$ by

$$(3.4) \qquad E = \left\{ (y, \delta) \,\middle|\, \begin{array}{l} y\epsilon\mathcal{Y},\ \delta\epsilon\mathfrak{m}_0, \\ \\ R_y(\delta, \theta) \leq R_y(\eta_{0,\,y}, \theta) \quad \text{for all} \quad \theta\epsilon\Theta \end{array} \right\}$$

Since $\mathfrak{m}_0$ is essentially complete, given $y\epsilon\mathcal{Y}$, there is a δ such that $(y, \delta)\epsilon E$. Hence $\text{proj}(E) = \mathcal{Y}$. The proof now proceeds in two steps:

(i) We first show that the Brown-Purves Theorem is applicable to E. Thus, there is a Borel measurable function $\rho: \mathcal{Y} \to \mathfrak{m}_0$ such that $(y, \rho(y))\epsilon E$ for all $y\epsilon\mathcal{Y}$.

(ii) Then it is shown that $\rho(y)$ can be "represented" by an element of $\mathfrak{m}^*$ (using the Farrell representation of continuous bilinear functionals).

Now, to the details. The space $\mathcal{Y}$ is a complete separable metric space by assumption and $\mathfrak{m}_0$ is a compact metric space so is complete and separable. Also

$$
\begin{aligned}
(3.5)\qquad E_y &= \{\delta\epsilon\mathfrak{m}_0 \mid R_y(\delta, \theta) \leq R_y(\eta_{0,\ y}, \theta) \text{ for all } \theta\epsilon\Theta\} \\
&= \bigcap_{\theta\epsilon\Theta} \{\delta \mid R_y(\delta, \theta) \leq R_y(\eta_{0,\ y}, \theta)\}.
\end{aligned}
$$

But $R_y(\cdot, \theta)$ is continuous on $\mathfrak{m}_0$ so E_y is compact for each y. Thus, to apply the Brown-Purves result, it remains to show that E is a Borel set. By (2.7), $R_y(\delta, \cdot)$ and $R_y(\eta_{0,\ y}, \cdot)$ are continuous on Θ. Let $\theta_1, \theta_2, \ldots$ be a countable dense set in Θ. Then it is easy to show that

$$(3.6)\qquad E = \bigcap_{i=1}^{\infty} E_{\theta_i}$$

where

$$(3.7) \qquad E_\theta = \{(y, \delta) | R_y(\delta, \theta) \leq R_y(\eta_{0, y}, \theta)\}.$$

Thus, to show E is Borel, it suffices to show that each E_θ is Borel. From Lemma A.1 in the Appendix, $R_y(\delta, \theta)$ is a Borel measurable function on $\mathcal{Y}\times\mathfrak{M}_0$ to R^1. Also, from Lemma A.2 in the Appendix $R_y(\eta_{0, y}, \theta)$ is Borel measurable on $\mathcal{Y}$ to R^1 so it is Borel measurable on $\mathcal{Y}\times\mathfrak{M}_0$ to R^1. Thus, $E_\theta = \{(y, \delta) | R_y(\delta, \theta) - R_y(\eta_{0, y}, \theta) \leq 0\}$ is a Borel set in $\mathcal{Y}\times\mathfrak{M}_0$.

Applying the Brown-Purves (1973) result, there is a Borel function $\rho : \mathcal{Y} \to \mathfrak{M}_0$ such that $(y, \rho(y)) \in E$ for all $y \in \mathcal{Y}$. Thus, for each $y \in \mathcal{Y}$

$$(3.8) \qquad R_y(\rho(y), \theta) \leq R_y(\eta_{0, y}, \theta) \quad \text{for all} \quad \theta \in \Theta.$$

To complete the proof, we now show that ρ corresponds to an element of $\mathfrak{M}^*$. Let $\tilde{\mathcal{L}}_1 \equiv L_1(\mathcal{X}\times\mathcal{Y}, \mathcal{B}_1\times\mathcal{B}_2, \mu\times\nu)$. For $h \in \tilde{\mathcal{L}}_1$, and $g \in \mathcal{C}(A)$, the function

$$(3.9) \qquad y \to \iint g(a)\rho_y(da|x)h(x, y)\mu(dx)$$

is Borel measurable and ν-integrable by Lemma A.3 in the Appendix. (Here, ρ_y is written for $\rho(y)$ for ease of reading.) Thus, the function $[\cdot, \cdot]$ on $\mathcal{C}(A)\times\tilde{\mathcal{L}}_1$ given by

$$(3.10) \qquad [g, h] = \iiint g(a)\rho_y(da|x)h(x, y)\mu(dx)\nu(dy)$$

is well defined. Clearly, $[\cdot, \cdot]$ is a bilinear functional on $\mathcal{C}(A)\times\tilde{\mathcal{L}}_1$. Further, it is easy to verify that $[\cdot, \cdot]$ satisfies (2.11) (i), (ii) and (iii) (with the subscript δ suppressed). Thus, by Farrell's (1964) result, there is a decision function $\eta:\mathcal{B}_A\times(\mathcal{X}\times\mathcal{Y}) \to [0, 1]$ such that

$$(3.11) \qquad [g, h] = \int\int\int g(a)\eta(da|x, y)h(x, y)\mu(dx)\nu(dy)$$

for $g\in\mathcal{C}(A)$ and $h\in\tilde{\mathcal{L}}_1$.

We now claim there is a ν-null set, say, N, such that if $y\notin N$, then $[\eta_y] = \rho(y)$. To see this, let $\{g_i\}_{i=1}^{\infty}$ and $\{f_j\}_{j=1}^{\infty}$ be countable dense sets in $\mathcal{C}(A)$ and L_1, respectively. For $h(x, y) = f_j(x)k(y)\in\tilde{\mathcal{L}}_1$, (3.10) and (3.11) yield

$$(3.12) \qquad \int[\int\int g_i(a)\rho_y(da|x)f_j(x)\mu(dx)]k(y)\nu(dy)$$

$$= \int[\int\int g_i(a)\eta(da|x, y)f_j(x)\mu(dx)]k(y)\nu(dy)$$

for all ν-integrable k. Thus there is a ν-null set $N_{i,j}$ such that if $y\notin N_{i,j}$,

$$(3.13) \qquad \int\int g_i(a)\rho_y(da|x)f_j(x)\mu(dx)$$

$$= \int\int g_i(a)\eta(da|x, y)f_j(x)\mu(dx).$$

Let $N = \bigcup_{i=1}^{\infty}\bigcup_{j=1}^{\infty} N_{i,j}$ so N is ν-null and if $y\notin N$, (3.13) holds for all i and j.

Hence if $y \notin N$, we have

$$(3.14) \qquad \iint g(a)\rho_y(da|x)f(x)\mu(dx)$$
$$= \iint g(a)\eta(da|x,\, y)f(x)\mu(dx)$$

for all $g \in C(A)$ and $f \in L_1$. Thus for $y \notin N$, $\rho(y) = [\eta_y]$. Since $\rho(y) \in \mathfrak{m}_0$ for all y, $\eta \in \mathfrak{m}^*$ by definition. This completes the proof.

Theorem 2: $\mathfrak{m}^*$ is an essentially complete class for the unconditional decision problem when the family of densities is

$$(3.15) \qquad \{f(x|y,\, \theta)k(y)\,|\,\theta \in \Theta,\ k \geq 0,\ \int k(y)\nu(dy) = 1\}.$$

Proof. This follows immediately from Theorem 1.

§4 Discussion and Applications

Suppose $U = (X,\, Y) \in \mathcal{R}^{p+1}$ is a random observable with an exponential family density of the form

$$(4.1) \qquad P_\eta(U \in B) = c(\eta)\int_B e^{\eta' u}\lambda(du)$$

where λ is a probability measure on $\mathcal{R}^{p+1}$. Partition η as $\eta' = (\theta,\, \omega')$ where $\theta \in \mathcal{R}^1$, $\omega \in \mathcal{R}^p$. The conditional density of X given $Y = y$ with respect to the conditional probability measure $\mu(dx|y)$ is

$$(4.2) \qquad f(x|y,\theta) = \frac{e^{\theta x}}{\int_{R^1} e^{\theta' z}\mu(dz|y)}.$$

Thus, the conditional on y, X has an exponential family distribution on R^1. Suppose we have a monotone multiple decision problem involving θ (see Ferguson (1967), Chapter 6 for the definition of monotone multiple decision problems and monotone decision rules). Then, according to Theorem 1 in Ferguson (1967)(Chapter 6), the class of monotone decision rules is essentially complete for each fixed y. To apply the results of Section 3, we assume that the family of measures $\{\mu(\cdot|y)|y \in R^p\}$ is dominated by a fixed σ-finite measure μ. The remaining assumptions necessary to apply Theorem 1 and Theorem 2 are easily checked. Thus the class of conditional (on y) monotone decision rules is essentially complete for the unconditional decision problem.

The assumption that the family of conditional measures $\{\mu(\cdot|y)|y \in R^p\}$ is dominated also must be made if one applies Theorems 1 and 2 to the testing problem treated in Matthes and Truax (1967). In the context of the Matthes-Truax (1967) paper, it is not hard to construct examples where the family of conditional probability measures is not dominated by a σ-finite measure. Thus, the results of the current work are not applicable to all of the testing problems to which the Matthes-Truax (1967) results or those in Eaton (1970) can be applied. However, the results established here are not restricted to testing problems nor to exponential families. For example, in a non-parametric context, Kariya and Eaton (1976) established a robustness property of the two-sided t-test using the Generalized Neyman-Pearson Lemma. Alternatively, this result may be established using Theorem 1 without recourse to the Generalized Neyman-Pearson Lemma.

Appendix

Throughout the appendix, the notation and assumptions of Section 1 hold.

Lemma A.1: Suppose $h: \mathfrak{X} \times \mathcal{Y} \to \mathcal{R}^1$ is $\mathcal{B}_1 \times \mathcal{B}_2$ measurable and $\int |h(x, y)| \mu(dx) < +\infty$ for each $y \in \mathcal{Y}$. For each $g \in \mathcal{C}(A)$, the function $T(y, \delta) \equiv \iint g(a)\delta(da|x)h(x, y)\mu(dx)$ on $\mathcal{Y} \times \mathfrak{M}$ to $\mathcal{R}^1$ is Borel measurable.

Proof: Let μ_0 be a probability measure on $(\mathfrak{X}, \mathcal{B}_1)$ with the same null sets as μ and let $f_0 = \frac{d\mu}{d\mu_0} \geq 0$. Then

$$T(y, \delta) = \iint g(a)\delta(da|x)h(x, y)f_0(x)\mu_0(dx).$$

Without loss of generality, $h \geq 0$ so $h(x, y)f_0(x)$ is the increasing limit of non-negative $\mathcal{B}_1 \times \mathcal{B}_2$ simple functions. Thus it suffices to show that

$$T_F(y, \delta) = \iint g(a)\delta(da|x)I_F(x, y)\mu_0(dx)$$

is Borel measurable where $F \in \mathcal{B}_1 \times \mathcal{B}_2$. However, the class of sets $\mathfrak{F} = \{F | T_F \text{ is Borel measurable}\}$ clearly contains all measurable rectangles, all disjoint unions of measurable rectangles, and complements of measurable rectangles. The Monotone Convergence Theorem shows that $\mathfrak{F}$ is a monotone class so $\mathfrak{F} = \mathcal{B}_1 \times \mathcal{B}_2$. This completes the proof.

Remark: The above proof is a minor modification of that in Sudderth (1971), (Section 5). A related reference in Dubins and Freedman (1964).

<u>Lemma A.2:</u> Let h be as in Lemma A.1, and suppose $\eta: \mathcal{B}_A \times (\mathcal{X} \times \mathcal{Y}) \to [0, 1]$ be a decision function. For each $g \in \mathcal{C}(A)$, the function

$$y \to \iint g(a)\eta(da|x, y)h(x, y)\mu(dx)$$

on $\mathcal{Y}$ to $\mathcal{R}^1$ is Borel measurable.

<u>Proof:</u> The proof is similar to that of Lemma A.1.

<u>Lemma A.3:</u> Let $\widetilde{L}_1 = L_1(\mathcal{X} \times \mathcal{Y}, \mathcal{B}_1 \times \mathcal{B}_2, \mu \times \nu)$. If $\rho: \mathcal{Y} \to \mathcal{M}$ is Borel measurable and if $h \in \widetilde{L}_1$ then the function

$$T(y) \equiv \iint g(a)\rho_y(da|x)h(x, y)\mu(dx)$$

is $\mathcal{B}_2$ measurable and ν-integrable.

<u>Proof:</u> If $h(x, y) = h_1(x)h_2(y)$ then $T(y) = h_2(y)T_1(y)$ where

$$T_1(y) = \int g(a)\rho_y(da|x)h_1(x)\mu(dx).$$

Since T_1 is the composition of the measurable map ρ and the continuous map T_{g, h_1} defined in (2.11), T_1 is measurable so $T = h_2 T_1$ is measurable. Now, arguing as in Lemma A.1, it follows that T is measurable. That T is ν-integrable is clear.

References

[1] Brown, L.D. and Purves, R. (1973). Measurable selections of extrema. Ann. Statist. 1, pp. 902-912.

[2] Dubins, L. and Freedman, D. (1964). Measurable sets of measures. Pac. Jour. of Math. 14, pp. 1211-1222.

[3] Eaton, M.L. (1970). A complete class theorem for multidimensional one-sided alternatives. Ann. Math. Statist. 41, pp. 1884-1888.

[4] Farrell, R.H. (1967). Weak limits of sequences of Bayes procedures in estimation theory. Proc. Fifth Berkeley Symp. Math. Statist. Prob. U. of California Press, 1, pp. 83-111.

[5] Ferguson, T. (1967). Mathematical Statistics, A Decision Theoretic Approach. Academic Press, New York.

[6] Kariya, T. and Eaton, M.L. (1976). Robust tests for spherical symmetry. Submitted to Ann. Statist.

[7] Matthes, T.K. and Truax, D.R. (1967). Tests of composite hypothesis for the multiparameter exponential family. Ann. Math. Statist. 38, pp. 681-697.

[8] Sudderth, W.D. (1971). On measurable gambling problems. Ann. Math. Statist. 42, pp. 260-269.